INFORMATION AND CONTROL 10, 475–484 (1967)

The Equivalence of Two Concepts of Categorial Grammar

JOEL M. COHEN

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts; and The University of Chicago, Chicago, Illinois

There are two types of categorial grammar systems, one studied by Bar-Hillel, Gaifman, and Shamir, and one studied by Lambek. They differ in the type of cancellation allowable. The former has only the rule $a(a\backslash b) \to b$ and $(a/b)b \to a$. The latter has an added rule that whenever $Xy \to z$ or $xY \to z$, then $X \to z/y$ or $Y \to x\backslash z$, respectively. The set of strings of words whose categories cancel to s (grammatical sentence) is called the language of the grammar being studied. This paper proves that a set of strings of words forms a categorial language of one type if and only if it forms a categorial language of the other type.

I. INTRODUCTION

Categorial grammar entails, among other requirements, the assignment of each word in a vocabulary to a class called a category. For example, in English we have the class of singular nouns, the class of intransitive verbs with plural subjects, the class of adjectives, etc. (This is not to say, however, that English will satisfy the other requirements.) There are certain rules of juxtaposition which will tell us whether or not a given string forms a grammatical sentence.

For example, we may assign to JOHN the singular noun category n, to RUNS the intransitive singular verb category $n\backslash s$ (meaning that when preceded by a category n, we get s = sentence). Then the string JOHN RUNS becomes in categorial terms $n(n\backslash s)$ which we say cancels to s. Thus JOHN RUNS is a sentence. The word KNOWS may be assigned to the singular intransitive-verb-with-singular-object category $(n\backslash s)/n$ so that JOHN KNOWS MARY becomes $n(n\backslash s)/n\ n \to n(n\backslash s) \to s$. Our basic rules, then, will be

$$a/b\ \ b \to a \quad \text{and} \quad a\ \ a\backslash b \to b. \tag{1.1}$$

In the last sentence JOHN KNOWS MARY we could have said that the verb category was $n\backslash(s/n)$ and we still would have had cancellation to s. Thus we could consider allowing a rule which says that we can always do this:

$$x\backslash(y/z) \rightarrow (x\backslash y)/z \quad \text{and} \quad (x\backslash y)/z \rightarrow x\backslash(y/z). \tag{1.2}$$

Since many words have different uses, we can allow each word in a vocabulary to be assigned to a finite number of categories. The word TIME could have the categories n, $n\backslash s/n$, and $n\backslash s/n^*$ (n^* is a plural noun).

A pronoun such as HE cannot be assigned to the category n, because it is only to replace a noun in the subject position of a sentence. Thus since $n(n\backslash s) \rightarrow s$, where n is used as a subject, we have the category for HE $s/(n\backslash s)$ so that $s/(n\backslash s)\ (n\backslash s) \rightarrow s$. Similarly for HIM we have $(s/n)\backslash s$. But then if we come to the string HE LIKES HIM, which clearly should be a sentence, we have $s/(n\backslash s)\ (n\backslash s)/n\ (s/n)\backslash s$, and no further reduction is possible. To take care of a case like this, we introduce the rule

$$x/y \quad y/z \quad \rightarrow \quad x/z. \tag{1.3}$$

Then our last sentence becomes $s/(n\backslash s)\ (n\backslash s)/n(s/n)\backslash s \rightarrow s/n\ (s/n)\backslash s \rightarrow s$.

One thing we would always want to be able to do in English is substitute JOHN for either HE or HIM in any sentence and still have a sentence. To insure that this is always the case we can use a rule allowing us to replace n by $s/(n\backslash s)$ or $(s/n)\backslash s$,

$$x \rightarrow y/(x\backslash y) \quad \text{and} \quad x \rightarrow (y/x)\backslash y. \tag{1.4}$$

We have now a set of four rules. It is not necessary that we consider a grammatical system using all of them. We may consider a grammatical system using simply rule (1.1). We will call a system that uses simply this one rule a *categorial grammar*, and a system which uses (1.1), (1.2), (1.3), and (1.4) a *free categorial grammar*.

II. CATEGORIAL GRAMMARS

We shall now formalize what has preceded.

(2.1) DEFINITION. Let $C' = \{x_1, \cdots, x_m\}$ be a finite set of symbols. A category system over C' is the infinite set C defined as follows:

(a) $C' \subset C$.

(b) If $x,y \in C$, then $x\backslash y$ and $x/y \in C$.

Notation. X, Y, Z, etc. represent strings of elements in C; x, y, z, etc. represent elements of C.

(2.2) DEFINITION. X directly cancels to Y, $X \Rightarrow^* Y$ if and only if

(i) $X = Y$ or

(ii) $X = W\, x/y\, yZ$ or $X = Wy\, y\backslash x\, Z$ and $Y = W\, x\ \ Z$.

X freely directly cancels to Y, $X \rightarrow* Y$, if and only if $X \Rightarrow* Y$ or

(iii) $X = (x\backslash y)/z$ and $Y = x\backslash(y/z)$ or vice versa.

(iv) $X = x/y\ y/z$ and $Y = x/z$ or $X = x\backslash y\ y\backslash z$ and $Y = x\backslash z$

(v) $X = x$ and $Y = y/(x\backslash y)$ or $Y = (y/x)\backslash y$

or

(vi) $X = WX'Z$ and $Y = WY'Z$ where $X' \rightarrow^* Y'$ by rules (i)–(v).

Then we can define (free) cancellation by letting $\Rightarrow$ (resp. $\rightarrow$) be the partial ordering relation generated by $\Rightarrow*$ (resp $\rightarrow^*$); i.e., $X \Rightarrow Y$ if there is a sequence $X_0, \cdots, X_n$ where $X_i \Rightarrow* X_{i+1}$ and $X_0 = X$, $X_n = Y$; similarly for $\rightarrow$.

(2.3) DEFINITION. A (free) categorial grammar, c.g. (f.c.g), is a quintuple (V, C, s, U, t), where V is a finite vocabulary, C a category system, s a distinguished element of C', U a function assigning to each $A \in V$ a finite subset $U(A) \subset C$, and $t = *$ (resp. $t = f$).

Let V^* be the set of finite strings of elements of V. Then $A_1 \cdots A_k \in V^*$ is said to be accepted by G if and only if there exist $y_i \in U(A_i)$, $i = 1, \cdots, k$ such that $y_1 \cdots y_k \Rightarrow s$ if $t = *$ and $y_1 \cdots y_k \rightarrow s$ if $t = f$. The set of strings in V^* accepted by G is the language of G, $L(G)$.

Two grammars (under any system) G and H are *weakly equivalent*, $G \simeq H$, if and only if $L(G) = L(H)$.

Free categorial grammar is what Lambek (1958) called categorial grammar; and what is here called categorial grammar, Bar-Hillel *et al.* (1960) called bidirectional categorial grammar.

We are going to prove that the two definitions are weakly equivalent; that is, any set of strings is a categorial language if and only if it is a free categorial language.

III. THE EQUIVALENCE OF f.c.g.'s AND c.g.'s

(3.1) THEOREM. *Given an f.c.g.G, there exists a c.g.H such that $G \simeq H$.*

Proof: According to Lambek (1958) we may replace the definition of $\rightarrow$ by

(a) $X \rightarrow X$.

(b) If $Y \to Y'$ then $(XY)Z \to X(Y'Z)$ and $X(YZ) \to (XY')Z$.
(c) If $X \to z/y$, then $Xy \to z$; if $Y \to x\backslash z$, then $xY \to z$.
(d) If $X \to Y$ and $Y \to Z$, then $X \to Z$.
(e) If $Xy \to z$, then $X \to z/y$; if $xY \to z$, then $Y \to x\backslash z$.

It is clear that (a), (b), (c), and (d) are defining relations for $\Rightarrow$. Thus the difference between $\Rightarrow$ and $\to$ is simply the application of rule (e). We shall show that by a certain modification of the assignment function U, rule (e) becomes unnecessary.

Consider the following rules:
(f) $x \to u/(x\backslash u)$ and $x \to (u/x)\backslash u$.
(g) $x/y \to (x/u)/(y/u)$ and $x\backslash y \to (u\backslash x)\backslash(u\backslash y)$.
(h) $x/y \to (u/x)\backslash(u/y)$ and $x\backslash y \to (x\backslash u)/(y\backslash u)$.
(k) $x \to y$ implies that $x/u \to y/u$ and $u\backslash x \to u\backslash y$.
(l) $x \to y$ implies that $u/y \to u/x$ and $y\backslash u \to x\backslash u$.

Let us call rules (a), (b), (c), and (d) rules I, and rules (f), (g), (h), (k), and (l) rules II.

(3.2) Lemma. *Under rules I, rule* (e) *is equivalent to rules II.*

Proof: See Appendix.

We assume that we are given the f.c.g. $G = (V, C, s, U, f)$. Let $D' = \bigcup_{A \in V} U(A)$. D' is finite. Then define D as follows: D is the smallest set containing D' and containing x and y whenever $x\backslash y$ or x/y is in D.

Definition. Consider the functions from C to C defined as follows: $e(x) = x$, $g_a(x) = x/a$, $h_a(x) = a\backslash x$, $k_a(x) = a/x$, $m_a(x) = x\backslash a$, for any $a \in C$. These are called the simple functions and we define $\sigma e = \sigma g_a = \sigma h_a = -\sigma k_a = -\sigma m_a = 1$. If $f_1, \cdots, f_n$ are simple functions we consider the composite $f_1 \circ f_2 \cdots \circ f_n = f$ and define $\sigma f = (\sigma f_1)(\sigma f_2) \cdots (\sigma f_n)$. Finally, we allow the constant function: for some a and for all x $f(x) = a$. $\sigma f = 1$. Any reference to function will always mean one of the above.

We now define $\bar{U}(A)$ for $A \in V$ as follows:
(1) $U(A) \subset \bar{U}(A)$
(2) If $\sigma f = 1$ and $f(x/y) \in \bar{U}(A)$ then for all $z \in D, f[(x/z)/(y/z)]$ and $f[(z/x)\backslash(z/y)] \in \bar{U}(A)$.
If $\sigma f = -1$ and $f[(x/z)/(y/z)]$ or $f[(z/x)\backslash(z/y)] \in \bar{U}(A)$ for any z, then $f(x/y) \in \bar{U}(A)$.
(3) If $\sigma f = 1$ and $f(x\backslash y) \in \bar{U}(A)$ then for all $z \in D$ $f[(z\backslash x)\backslash(z\backslash y)]$ and $f[(x\backslash z)/(y\backslash z)] \in \bar{U}(A)$. If $\sigma f = -1$ and $f[(z\backslash x)\backslash(z\backslash y)]$ or $f[(x\backslash z)/(y\backslash z)] \in \bar{U}(A)$ for any z, then $f(x\backslash y) \in \bar{U}(A)$.

(4) If $\sigma f = 1$ and $f(x) \in U(A)$, then for all $y \in D$, $f[(y/x)\backslash y]$ and $f[y/(x\backslash y)] \in \bar{U}(A)$. If $\sigma f = -1$ and for any $y \in D$ $f[(y/x)\backslash y]$ or $f[y/(x\backslash y)] \in U(A)$, then $f(x) \in \bar{U}(A)$.

The following properties of $\bar{U}$ are immediate:

(a) $\bar{U}(A)$ is finite since D is finite.

(b) If $y \in \bar{U}(A)$, then there is some $x \in U(A)$ such that $x \to y$.

From (a) we see that we can define a categorial grammar $H = (V, C, s, \bar{U}, *)$. If $A_1 \cdots A_n \in L(H)$, then there exist $y_i \in \bar{U}(A_i)$ such that $y_1 \cdots y_n \Rightarrow s$. But from (b) there exist $x_i \in U(A_i)$ such that $x_i \to y_i$, hence $x_1 \cdots x_n \to y_1 \cdots y_n \to s$. Thus $A_1 \cdots A_n \in L(G)$.

Thus $L(H) \subset L(G)$. We claim that $L(G) \subset L(H)$. Thus we must prove that if there exist $x_i \in U(A_i)$ such that $x_1 \cdots x_n \to s$, then there exist $y_i \in \bar{U}(A_i)$ such that $y_1 \cdots y_n \Rightarrow s$.

Since $\to$ and $\Rightarrow$ differ only by the use of rules II [using Lemma (3.2)] we must show that any application of rules II in reducing $x_1 \cdots x_n$ to s may be omitted by replacing some $x_i \in U(A_i)$ by some $y_i \in \bar{U}(A_i)$.

Observe first that we can eliminate application of the rules (f), (g), or (h) for any $u \in D$. Say we used rule (f) to get $WxZ \to W(u/x)\backslash uZ$ for $u \in D$. Clearly in order to have cancelled to WxZ we must have had some $x_i = f(x)$ where $\sigma f = 1$. But then $f[(u/x)\backslash u] = y_i \in \bar{U}(A_i)$. Thus $x_1 \cdots x_{i-1} y_i x_{i+1} \cdots x_n$ cancels to $W(u/x)\backslash uZ$ with one fewer applications of rule (f) than $x_1 \cdots x_n$ requires. Similarly for the other part of (f) and for (g) and (h).

We shall next show that applications of (f), (g), and (h) for u outside D are unnecessary. Then we have that when $x \to y$ is necessary we may *replace* $f(x)$, wherever it appears and $\sigma f = 1$, by $f(y)$ and we may replace $f(y)$, wherever it appears and $\sigma f = -1$, by $f(x)$. Thus (k) and (l) "hold," so $\to s$ and $\Rightarrow s$ are equivalent using $\bar{U}$ and Theorem (3.1) is proved.

Assume that rule (f) was used in this situation: $WxZ \to s$ by using the fact that $WxZ \to Wy/(x\backslash y)Z$. We also assume that this is the furthest left application of any of rules II (i.e., no *free* cancellation takes place in W). We assume further that $y \notin D$, thus y cannot appear in W. We assume further that all cancellation involving Z alone has been done. There are three cases, one of which must occur in order for this string to cancel to s:

(1) $Z = x\backslash yZ'$ where $WyZ' \to s$;

(2) $Z = (x\backslash y)/wZ'$ where $Wy/wZ' \to s$;

(3) $Z = (y/(x\backslash y))\backslash wZ'$ where $WwZ' \to s$.

We shall now see what happens in these three cases.
(1) $WxZ = Wx\ x\backslash yZ' \Rightarrow WyZ'$ thus we have gotten to the point we wanted and have eliminated one application of rule (f).
(2) Since $y \notin D$, $(x\backslash y)/w \notin D$, thus $(x\backslash y)/w \notin D'$, thus $(x\backslash y)/w \notin U(A)$ for any $A \in V$. Thus $u = (x\backslash y)/w$ must have been gotten from some application of rules II [or their equivalent, rule (e)]. That is, we must have had $u' \rightarrow u$ for some $u' \neq u$ using the fact that $u'w \rightarrow x\backslash y$ implies that $u' \rightarrow (x\backslash y)/w = u$. Then since $u'w$ must involve y, this must in turn have come from the fact that $xu'w \rightarrow y$ implies $u'w \rightarrow x\backslash y$. But then we could also have had $xu' \rightarrow y/w$ and then $u' \rightarrow x\backslash(y/w)$. Then using the *same number* of previous applications of rule (e), we could have had $x\backslash(y/w)$ instead of $(x\backslash y)/w$. But then we could have had $Wx\ x\backslash(y/w)Z'$ instead of $Wx\ (x\backslash y)/wZ'$. But $Wx\ x\backslash(y/w)Z' \Rightarrow W(y/w)Z'$ so we have gotten to the point we wanted and have eliminated one application of rule (e).
(3) By the same reasoning as at the beginning of case (2), $u = [y/(x\backslash y)]\backslash w$ must have come from some application of rule (e) from some element $u' \neq u$. Then $[y/(x\backslash y)]u' \rightarrow w$. Since $u' \neq u$, there are two choices for u'. $u' = x\backslash y$ and $w = y$; or $u' = (x\backslash y)/z$ and $w = y/z$. In the former case we could have gone directly from $Wxu'Z' = Wx\ x\backslash yZ' \Rightarrow WyZ' = WwZ'$ to the place where we wished to be with the application of rule (e) eliminated. In the latter case, an argument similar to that of case (2) shows that we could have had in place of $(x\backslash y)/z = u_1'\ x\backslash(y/z) = u''$; then $Wxu''Z' = Wx\ x\backslash(y/z)Z' \Rightarrow Wy/zZ' = WwZ'$, which is where we wanted to be having eliminated one application of rule (e).

[*Note:* There is an alternative argument to case (2) which may be easier to follow intuitively. In order for further reduction to occur, we must have one of the following three subcases: $Z' = wZ''$, $Z' = w/uZ''$, or $Z' = [(x\backslash y)/w]\backslash uZ''$. In the first case, everything works out easily: $WxZ = Wx(x\backslash y)/w\ wZ'' \Rightarrow WyZ''$ which is what we want. The third case reduces either to the first or the second (as in case (3) above). Then for further cancellation in the second subcase we have three sub-sub-cases: $Z'' = uZ'''$, $Z'' = u/vZ'''$, or $Z'' = (w/u)\backslash vZ'''$, and the whole process begins again. Since Z, Z', Z'', Z''' are all of finite length the case (2)'s cannot continue indefinitely, so eventually we must come to a case (1) and the problem is solved. We have eliminated the single application of rule (e).]

The other part of rule (f) can be handled analogously. That leaves (g) and (h), which are similar enough to choose one-half of one rule.

Assume $Wx/yZ \to s$ by using the fact that $Wx/yZ \to W(x/u)/(y/u)Z$ for $u \notin D$. Again assume that this is the farthest left application of rules II. We have three cases:

(1) $Z = y/uZ'$ where $Wx/uZ' \to s$.

(2) $Z = (y/u)/wZ'$ where $W(x/u)/wZ' \to s$.

(3) $Z = ((x/u)/(y/u))\backslash wZ'$ where $WwZ' \to s$.

Case (1) has three subcases: $Z' = uZ''$ where $WxZ'' \to s$; $Z' = u/wZ''$; and $Z' = (y/u)\backslash wZ''$. Again the very first subcase is easily shown to make rule (g) unnecessary: $Wx/yZ = Wx/y\ y/u\ u\ Z'' \Rightarrow Wx/y\ yZ'' \Rightarrow WxZ''$, which is where we wanted to be having eliminated one occurrence of rules II. Subcases (2) and (3) and cases (2) and (3) will be similar to the previous and will all eventually reduce to case (1) subcase (1) as in the proof of rule (1).

Thus we have completed the proof of this theorem. H is weakly equivalent to G.

We now prove the converse of Theorem (3.1):

(3.3) THEOREM. *Given a c.g. G, there exists an f.c.g. H such that $H \simeq G$.*

Proof: For this theorem we must introduce a special type of category system.

(3.4) DEFINITION. Let C' be as in (2.1). A *restricted category system* over C' is a finite set $\bar{C} = C' \cup \{x/y\colon x, y \in C'\} \cup \{(x/y)/z\colon x, y, z \in C'\}$.

A *restricted category grammar*, r.c.g., $G = (V, C, s, U, *)$ is a categorial grammar such that for all $A \in V$, $U(A) \subset \bar{C}$.

It has been proved by Bar-Hillel *et al.* (1960) that, given any c.g. G, there is an r.c.g. G' weakly equivalent to G. If G is our original c.g., let $G' = (V, C, s, U, *)$ be an r.c.g. weakly equivalent to G. Let $H = (V, C, s, U, f)$ be the *free* c.g. whose components are the same as those of G'. We shall show that H is weakly equivalent to G', and hence to G. What we shall show is that in cancelling a string in H to s, we would never have to make use of rules II.

Assume we have a string which reduces to s by using, a certain point, rule (f). That is, $WxZ \to Wy/(x\backslash y)Z$. Let us assume further that this is the farthest left application of any of rules II. Then since $y/(x\backslash y)$ and $(x\backslash y)$ are not in $\bar{C}$, they are not in $U(A)$ for any $A \in V$. Thus the argument used in the proof of (3.1) holds here, and we can eliminate the need for rule (f).

Similarly we can use this argument to eliminate the need for applying rules (h) and (k), and the second part of (g). That is because everything which occurs in these rules necessarily makes use of categories outside $\bar{C}$. The only remaining part of rules II is the application of rule (g) to the situation $Wx/y\ y/zZ \rightarrow Wx/zZ$. That is because x/y, y/z, and x/z may all be in $\bar{C}$. We may assume that W and Z contain no applications of rules II other than this use of rule (g). But then for further cancellation to occur, we must have $Z = (z/z_1)(z_1/z_2) \cdots (z_n/z_{n+1})z'_{n+1} Z'$ where $WxZ' \rightarrow s$. But then $W(z/y)(y/z)Z' = W(x/y)(y/z)(z/z_1)(z_1/z_2) \cdots (z_n/z_{n+1})z_{n+1}Z' \Rightarrow W(x/y)(y/z) \cdots z_nZ \Rightarrow W(x/y)yZ' \Rightarrow WxZ'$. Thus we have gotten to this same place without applying rule (g).

Thus rules II are unnecessary if $U(A) \subset \bar{C}$ for all $A \in V$. Then we have H weakly equivalent to G' and to G. So the theorem is proved.

Acknowledgment

The author wishes to thank Professor Noam Chomsky of the Massachusetts Institute of Technology for his encouragement and his helpful suggestions. This research was partially supported by NSF grant GP-5609.

REFERENCES

Bar-Hillel, Y., Gaifman, C., and Shamir, E. (1960). On categorial and phrase-structure grammars. *Bull. Res. Council Israel* **9F**, 1–16.

Lambek, J. (1958). The mathematics of sentence structure. *Am. Math. Monthly* **65**, 154–170.

Lambek, J. (1959). Contributions to a mathematical analysis of the English verb-phrase. *J. Can. Linguistic Assoc.* **5**, 83–89.

IV. APPENDIX

Proof of (3.2): Look at the rule

(m) If $Xy \rightarrow z$, then $X\ y/w \rightarrow z/w$; and if $xY \rightarrow z$, then $(w\backslash x)Y \rightarrow w\backslash z$.

First we shall show that rules I and II imply (m) and then use (m) to imply (e). Finally we shall show that rules I and (e) imply rules II.

Assume $Xy \rightarrow z$. Then after carrying out all the necessary cancellation involving only members of X we get a string $X' = x_1 \cdots x_n$ where $X \rightarrow X'$ and after carrying out transformations on y we get a term y^0 where $y \rightarrow y^0$ and there are four cases, one of which must be true:

(1) $x_n = x_n'/y^0$ where $x_1 \cdots x_{n-1}x_n' \rightarrow z$.

(2) $x_n = x_n'/y'$ where $y^0 = y'/u$ and $x_1 \cdots x_{n-1}(x_n'/u) \rightarrow z$.

(3) $y^0 = x_n\backslash y'$ where $x_1 \cdots x_{n-1}y' \rightarrow z$.

(4) $y^0 = x_n'\backslash y'$ where $x_n = x_n''\backslash x_n'$ and $x_1 \cdots x_{n-1}(x_n''\backslash y') \rightarrow z$.

We shall prove that $Xy/w \to z/w$ by induction on n, the second half of (m), holding by analogy.

For $n = 0$ X is empty and $y \to z$, so using rule (k) $y/w \to z/w$. We assume (m) holds up to $n - 1$. Now we separate into the four cases above.

(1) $X(y/w) \to X'(y/w) \xrightarrow{(k)} X'(y^0/w) = x_1 \cdots x_{n-1}(x_n'/y^0)(y^0/w)$

$$\xrightarrow{(g)} x_1 \cdots x_{n-1}(x_n'/w) \to z/w \text{ by induction.}$$

(2) Since $x_1 \cdots x_{n-1}(x_n'/u) \to z$, we have by induction

$$x_1 \cdots x_{n-1}[(x_n'/u)/w] \to z/w. \text{ But then}$$

$$X(y/w) \to x_1 \cdots x_{n-1}(x_n'/y')(y^0/w) = x_1 \cdots x_{n-1}(x_n'/y')[(y'/u)/w]$$

$$\xrightarrow{(y)} x_1 \cdots x_{n-1}[((x_n'/u)/w)/((y'/u)/w)]((y'/u)/w)$$

$$\to x_1 \cdots x_{n-1}[(x_n'/u)/w] \to z/w.$$

Cases (3) and (4) may be handled similarly. Thus we are now armed with rule (m) in proving (e).

We assume that $Xy \to z$ and wish to prove that $X \to z/y$ (the other part of (e) may be proved by analogy). We have $X \to x_1 \cdots x_n$ and $y \to y_0$ where $n \geqq 1$ (if $n = 0$, X is empty and it would never be necessary to use the rule $\emptyset \to z/y$ in order to get cancellation to s).

There are four cases for $n = 1$, one of which must be true if x_1 and y^0 are sufficiently reduced:

(1′) $x_1 = z/y^0$ thus $X \to x_1 = z/y^0 \xrightarrow{(l)} z/y$ since $y \to y_0$.
(2′) $x_1 = x'/y'$ and $y^0 = y'/u$ and $x'/u \to z$. Thus

$$X \to x_1 = x'/y' \xrightarrow{(g)} (x'/u)/(y'/u) = (x'/u)/y^0 \xrightarrow{(k)} z/y^0 \xrightarrow{(l)} z/y.$$

(3′) $y^0 = x_1\backslash z$. Thus $X \to x_1 \xrightarrow{(f)} z/(x_1\backslash z) = z/y^0 \xrightarrow{(l)} z/y$.
(4′) $x_1 = x'\backslash y'$, $y^0 = y'\backslash u$ and $x'\backslash u \to z$. Thus

$$X \to x_1 = x'\backslash y' \xrightarrow{(g)} (x'\backslash u)/(y'\backslash u) = (x'\backslash u)/y^0 \xrightarrow{(k)} z/y^0 \xrightarrow{(l)} z/y.$$

Thus (e) holds if $n = 1$. Assume it holds for $n - 1$. Assume that $X \to x_1 \cdots x_n$, $y \to y_0$ have been reduced as much as possible by themselves and $Xy \to z$. Then we have four cases as in the proof of (m), and these are handled as follows:

(1) $X \to x_1 \cdots x_{n-1}(x_n'/y^0) \xrightarrow{(m)} z/y_0 \xrightarrow{(l)} z/y.$

(2) $X \to x_1 \cdots x_{n-1}(x_n'/y') \xrightarrow{(g)} x_1 \cdots x_{n-1}((x_n'/u)/(y'/u))$
$\xrightarrow{(m)} z/(y'/u) = z/y^0 \xrightarrow{(l)} z/y.$

(3) $X \to x_1 \cdots x_n \xrightarrow{(f)} x_1 \cdots x_{n-1}(y'/(x_n\backslash y')) \xrightarrow{(m)} z/(x_n\backslash y')$
$= z/y^0 \xrightarrow{(l)} z/y.$

(4) $X \to x_1 \cdots x_{n-1}(x_n''\backslash x_n') \xrightarrow{(h)} x_1 \cdots x_{n-1}((x_n''\backslash y')/(x_n'\backslash y'))$
$\xrightarrow{(m)} z/(x_n'\backslash y') = z/y^0 \xrightarrow{(l)} z/y.$

Thus by induction rule (e) holds. Thus rules I and II imply (e). We now show that rules I and rule (e) imply rules II. Assuming rules I and rule (e), we have

(f): Since $x(x/y) \to y$, then by (e) $x \to y/(x\backslash y)$. Since $(y/x)x \to y$, then we have $x \to (y/x)\backslash y$.

(g)(h): Since $(x/y)(y/u)u \to (x/y)y \to x$, we have $(x/y)(y/u) \to x/u$, so $x/y \to (x/u)/(y/u)$ and $(y/u) \to (x/y)\backslash(x/u)$. Since $x(x\backslash y)(y\backslash u) \to u$, $(x\backslash y)(y\backslash u) \to x\backslash u$, thus we have $x\backslash y \to (x\backslash u)/(y/u)$ and $y\backslash u \to (x\backslash y)\backslash(x\backslash u)$.

(k): If $y \to z$, then $(y/w)w \to y \to z$ so that $(y/w) \to z/w$. If $y \to z$, then $w(w\backslash y) \to z$ so that $(w\backslash y) \to w\backslash z$.

(l): If $x \to y$ then $(z/y)x \to (z/y)y \to z$ and $x(y\backslash z) \to y(y\backslash z) \to z$ so $z/y \to z/x$ and $y\backslash z \to x\backslash z$.

Thus rules I and (e) imply rules II. So that under rules I, (e) is equivalent to rules II, and (3.2) is proved.

RECEIVED: October 13, 1965; revised April 14, 1967